Schedae Informaticae Vol. 22 (2013): 19–25
doi: 10.4467/20838476SI.13.002.2087

On Sequence Entropy of Thue-Morse Shift

Magdalena Foryś
Institute of Computer Science, Jagiellonian University,
Prof. Stanisława Łojasiewicza 6, 30-348 Kraków, Poland
e-mail: *magdalena.forys@uj.edu.pl*

Abstract. The paper summarizes properties of topological and sequence entropy of the Morse shift $X_{\mathcal{M}}$ generated by the Thue-Morse sequence $t_{\mathcal{M}}$. The first part is an estimation of growth rate of possible subwords in $t_{\mathcal{M}}$. We show a polynomial upper bound on the number of finite subwords occuring in $t_{\mathcal{M}}$ which is $Cn^{2\log 3}$ for some constant $C > 0$. In the second part we prove that the sequence entropy of $X_{\mathcal{M}}$ is achieved for the sequence $\tau(i) = 2^{2i} - 1$.

Keywords: entropy, Thue-Morse sequence, sequence entropy, pattern, pattern complexity.

1. Basic notions and definitions

Let $\mathcal{A}$ be a two-element set $\{0, 1\}$, and $\mathcal{A}^*$ denote a free monoid generated by $\mathcal{A}$ together with the operation of concatenation, defined for any $a = a_0 \dots a_m$, $b = b_0, \dots, b_n \in \mathcal{A}^*$ by the following formula:

$$ab = a_0 \dots a_m b_0 \dots b_m.$$

By ε we denote the empty word, which is the neutral element for concatenation. We say, that $\mathcal{A}$ is an alphabet, and the elements of the free monoid generated by this alphabet are words. Let us consider a set of infinite sequences over the alphabet $\mathcal{A}$:

$$\mathcal{A}^{\mathbb{N}} = \{x = (x_n)_{n\in\mathbb{N}} : x_n \in \mathcal{A} \text{ for all } n \in \mathbb{N}\}.$$

For $x \in \mathcal{A}^{\mathbb{N}}$ every finite subsequence of x is called a subword. For every finite word $x = x_0 \dots x_n$ we may define a word $\overline{x}$ which arises from x by changing every 0 into 1

and every 1 into 0. Such a word is called a complement of x. The length of a word $x = x_0 \dots x_n$ is a number of all letters which occur in it and is denoted by $|x|$. We define a mapping $\sigma : \mathcal{A}^{\mathbb{N}} \to \mathcal{A}^{\mathbb{N}}$ as follows:

$$(\sigma(x))_n = x_{n+1} \text{ for all } n \in \mathbb{N}.$$

The mapping σ is called a shift mapping or simply a shift. For any infinite sequence $x \in \mathcal{A}^{\mathbb{N}}$ we define an orbit of x:

$$\mathcal{O}(x) = \{\sigma^n(x) \in \mathcal{A}^{\mathbb{N}} : n \in \mathbb{N}\}.$$

$\mathcal{A}^{\mathbb{N}}$ together with a shift σ is a topological space, where the topology is given by the metric:

$$d(x,y) := \begin{cases} 2^{-\min\{n\in\mathbb{N}:\ x_n \neq y_n\}} & \text{for } x \neq y, \\ 0 & \text{otherwise} \end{cases}$$

Definition 1 *$X \subset \mathcal{A}^{\mathbb{N}}$ is called a shift space (or equivalently a shift) iff*

1. *X is a closed set,*
2. *X is σ-invariant, which means $\sigma(X) \subset X$.*

Given an infinite sequence $x \in \mathcal{A}^{\mathbb{N}}$ we may define a shift generated by this element $X_x = \overline{\mathcal{O}(x)}$. Such a construction assures that both conditions from the shift's definition are fulfilled.

2. Thue-Morse sequence

We consider the Thue-Morse sequence $t_{\mathcal{M}}$ over the alphabet $\{0,1\}$. Let us recall two equivalent definitions of this sequence:

Definition 2 *Thue-Morse sequence is defined by following formula:*

$$t_{\mathcal{M}} = \lim_{n\to\infty} \mu^n(0),$$

where $\mu : \{0,1\} \to \{0,1\}^*$ is a substitution such that:

$$\mu(0) = 01, \qquad \mu(1) = 10.$$

Remark 3 *Another way is to define the sequence $t_{\mathcal{M}}$ recurrently by the following formula:*

$$t_{\mathcal{M}} = \lim_{n\to\infty} B_n,$$

where for all $n \in \mathbb{N} : B_n \in \{0,1\}^*$ such that:

$$B_0 = 0, \qquad B_{n+1} = B_n\overline{B_n}.$$

where $\overline{B_n}$ is the complement of B_n.

Definition 4

1. *The sequence $x \in \mathcal{A}^{\mathbb{N}}$ is minimal iff every finite subword w of x occurs in x infinitely many times and the length of gaps between those occurences is bounded.*
2. *The shift $X \subset \mathcal{A}^{\mathbb{N}}$ is minimal iff for every $x \in X$ there is $\overline{\mathcal{O}(x)} = X$.*

The following fact, proved in [5], presents relations between the minimality of the sequence and of the shift generated by that sequence.

Fact 5 *For any minimal $x \in \mathcal{A}^{\mathbb{N}}$ the shift $X_x = \overline{\mathcal{O}(x)}$ is minimal.*

It can be proved that the Thue-Morse sequence is minimal and so is the Morse shift $X_{t_{\mathcal{M}}} = \overline{\mathcal{O}(t_{\mathcal{M}})}$. In the sequel we denote $X_{\mathcal{M}} = X_{t_{\mathcal{M}}}$

3. Topological entropy of the Morse shift

The main object of our consideration in this paper is the entropy of the Morse shift $X_{\mathcal{M}}$ generated by the Thue-Morse sequence. Let $\mathcal{B}_n(x)$ denote the set of all subwords of length n occuring in the infinite sequence x. In general case for a two-element-alphabet there are at most 2^n different words of the length n, so $\sharp\mathcal{B}_n(x) \leq 2^n$. If every word of the length n has the same probability of occurence in the sequence (equal to 2^{-n}) then we have a uniform distribution. Entropy tells us how much the actual distribution of words differs from the uniform one.

In the sequel we assume that writing log we mean the function $\log_2$. The following definition of entropy of a shift is a consequence of the fact, that for a minimal sequence x all of its subwords are at the same time all possible subwords occuring in elements of X_x.

Definition 6 *The (topological) entropy of a shift X_x generated by the sequence $x \in \mathcal{A}^{\mathbb{N}}$ is defined by the formula:*

$$h(X_x) = \lim_{n\to\infty} \frac{1}{n} \log \sharp\mathcal{B}_n(x).$$

The definition implies that $h(X_x) \in [0, 1]$ for every sequence $x \in \mathcal{A}^{\mathbb{N}}$.

Let us concentrate on the substitution μ which defines the sequence $t_{\mathcal{M}}$. It assures that it is possible to divide $t_{\mathcal{M}}$ into blocks of the form $01, 10$ of the length 2, starting from the beginning of the sequence. Analogously if we take any finite subword of $t_{\mathcal{M}}$ we may suspect that the same structure can be found inside it. The only positions which can possibly disturb that structure are the beginning and the end of a subword, where the remaining parts may be too short to fit in the schema.

Such a reasoning allows us to find some upper bound on number of words occuring in $t_{\mathcal{M}}$. It has its source in [4] and it eventually let us prove what the exact value of topological entropy for $t_{\mathcal{M}}$ is.

Lemma 7 *Let ω be a subword of $t_{\mathcal{M}}$ such that $|\omega| \geq 7$. Then there exists the unique decomposition:*

$$\omega = lur,$$

where $u \in \{01, 10\}^*$, $l, r \in \{\varepsilon, 0, 1\}$.

Proof. The sequence $t_{\mathcal{M}}$ can be divided into blocks of the form $01, 10$. If we start from the position with index 0, then every block of such a form is in the even position $t_{2n}t_{2n+1}$ for some $n \in \mathbb{N}$. Pairs 00 and 11 can occur between these blocks only. It is possible to divide the sequence into blocks $0110, 1001$ of length 4 starting from the beginning of the sequence.

Now, if a word ω contains only one block 11 or 00, then it is placed in the middle of some block of length 4 – respectively 0110 and 1001. In that case decomposition into blocks 01 and 10 is the unique indeed.

If there are more blocks 11 or 00 then the decomposition of $t_{\mathcal{M}}$ can be assigned to ω which is a subword of the sequence and that decomposition is unique as well. If there is a rest, then it is of length at most 1 and the thesis is satisfied for $l, r \in \{\varepsilon, 0, 1\}$.

The middle word u from the above lemma is a subword of $t_{\mathcal{M}}$ built from blocks $01, 10$, so there exists a word $v \in \{0, 1\}^*$ such that $u = \mu(v)$. In that case, the decomposition from the above lemma takes form:

$$\omega = l\mu(v)r,$$

for some $l, r \in \{\varepsilon, 0, 1\}$ and $v \in \{0, 1\}^*$. That observation implies the following lemma:

Lemma 8 *Let ω be a subword of the Thue-Morse sequence $t_{\mathcal{M}}$ such that $|\omega| \geq 7$. There exists the unique decomposition:*

$$\omega = l_0 \dots l_{k-1}\mu^k(u)r_{k-1} \dots r_0,$$

such that $l_i, r_i \in \{\varepsilon, \mu^i(0), \mu^i(1)\}$, $u \in \{0, 1\}^*$, $|u| \leq 6$.

Proof. From the previous lemma it follows that there exists a decomposition $\omega = l_0\mu(u_0)r_0$. Observe that there exists some $n \in \mathbb{N}$ such that ω is a subword of $B_n = \mu(B_{n-1})$, where B_n is a block from the reccurent definition of $t_{\mathcal{M}}$. Hence $\mu(u_0)$ is a subword of $\mu(B_{n-1})$, so u_0 is a proper subword of $t_{\mathcal{M}}$. If $|u_0| \geq 7$ we can apply the lemma again to get the decomposition:

$$\omega = l_0\mu(u_0)r_0 = l_0\mu(l_1'\mu(u_1)r_1')r_0 = l_0l_1\mu^2(u_1)r_1r_0.$$

We can repeat this reasoning as long as $|u_k| \geq 7$.

To calculate the entropy of $t_{\mathcal{M}}$ we need to find some upper bound on the number of subwords of length n which occur in the Thue-Morse sequence. The following theorem tells us, that this number can be estimated polynomially.

Theorem 9 *There exists a constant $C > 0$ such that for all $n \in \mathbb{N}$ we have:*

$$\sharp\mathcal{B}_n(t_{\mathcal{M}}) \leq Cn^{2\log 3}.$$

Proof. Let us fix $n \geq 0$. The above lemmas imply that the following bound on the number of elements of this set is true:

$$\sharp\mathcal{B}_n(t_{\mathcal{M}}) \leq \sharp\{\omega = l_0 \dots l_{k-1}\mu^k(u)r_{k-1} \dots r_0 : l_i, r_i \in \{\varepsilon, \mu^i(0), \mu^i(1)\}, |u| \leq 6\}.$$

Let us notice that l_i and r_i can take one from three possible values. There are also a finite number of values which a word u can take and let α denote that number. It is also true that the length of u is greater than 3. Another useful observation is that for a word of the length n the number k is always smaller than $\log n$.

The upper bound for the number of words of length n in the Thue-Morse sequence would be as follows:

$$\sharp\mathcal{B}_n(t_{\mathcal{M}}) < 2\alpha \cdot 3^{2\log n} = 2\alpha \cdot n^{2\log 3},$$

if only we are able to show that the number k take one from at most two possible values.

Let us notice that for every $i = 0, \dots, k-1$:

$$0 \leq |l_i| \leq 2^i \quad \text{and} \quad 0 \leq |r_i| \leq 2^i.$$

As we said earlier it is true that $3 \leq |u| \leq 6$, so we have:

$$3 \cdot 2^k \leq |\mu^k(u)| \leq 6 \cdot 2^k.$$

Since for an arbitrary word $\omega = l_0 \dots l_{k-1}\mu^k(u)r_{k-1} \dots r_0$ of length n we have:

$$3 \cdot 2^k \leq n \leq 2 \cdot \sum_{i=0}^{k-1} 2^i + 6 \cdot 2^k$$

and therefore

$$2^{k+1} < n < 7 \cdot 2^k < 8 \cdot 2^k = 2^{k+3}.$$

Then:

$$\log n - 3 < k < \log n - 1.$$

As k is a natural number it can take at most

$$\sharp[(\log n - 3, \log n - 1) \cap \mathbb{N}] \leq 2$$

values indeed. With such an upper bound we can now easily prove the following remark about the exact value of entropy of the Thue-Morse sequence.

Remark 10 *The entropy of the Morse shift $X_{t_{\mathcal{M}}}$ is equal to zero.*

Proof.

$$h(t_{\mathcal{M}}) = \lim_{n\to\infty} \frac{1}{n} \log \sharp\mathcal{B}_n(t_{\mathcal{M}}) \leq \lim_{n\to\infty} \frac{1}{n} \log 2\alpha n^{2\log 3} = 0.$$

4. Sequence entropy

In this section we consider the sequence entropy of the Morse shift. Let us take an increasing sequence $\tau \in \mathbb{N}^{\mathbb{N}}$ and a sequence $x \in \mathcal{A}^{\mathbb{N}}$. Let us fix some $n \in \mathbb{N}$ and for all $k \in \mathbb{N}$ let $x^{(n)}_{[k+\tau]}$ denote:

$$x^{(n)}_{[k+\tau]} = x_{k+\tau(0)} x_{k+\tau(1)} \dots x_{k+\tau(n-1)}.$$

If n is clear from the context we simply write $x_{[k+\tau]}$.

Definition 11 *A word w of length n is called an n-pattern of a sequence x according to the sequence τ if there exists some $k \in \mathbb{N}$ such that $x^{(n)}_{[k+\tau]} = w$.*

Definition 12

1. *The pattern complexity $p_x(n, \tau)$ is a number of different n-patterns occuring in sequence x according to the sequence τ.*
2. *The maximal pattern complexity $p^*_x(n) = \sup_\tau p_x(n, \tau)$, where the supremum is taken over all increasing sequences $\tau \in \mathbb{N}^{\mathbb{N}}$.*

The following fact is true:

Fact 13 *The maximal pattern complexity for t_{M} equals 2^n for $n = 1, 2, \dots$*

Proof of the above fact can be found in [2].

Definition 14 *Let (X, σ) be a shift, $n = 1, 2, \dots$ and $\epsilon > 0$.*

1. *The set $W \subset X$ (τ, ϵ, n)-spans some $B \subset X$ iff:*

$$\forall x \in B \; \exists y \in W : \; d(\sigma^{\tau(i)}(x), \sigma^{\tau(i)}(y)) < \epsilon \; \forall i = 1, \dots, n.$$

2. *The set $W \subset X$ is (τ, ϵ, n)-spanning iff it (τ, ϵ, n)-spans X.*

By $Span(\tau, \epsilon, n)$ we denote the smallest cardinality of all (τ, ϵ, n)-spanning sets.

Definition 15 *The sequence entropy along the sequence τ for a shift X is defined by the following formula:*

$$h_\tau(X) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log Span(\tau, \epsilon, n).$$

We define the sequence entropy of a shift X by the formula:

$$h_\infty(X) = \sup_\tau h_\tau(X),$$

where the supremum is taken again over all increasing sequences $\tau \in \mathbb{N}^{\mathbb{N}}$.

The first question that arises from the above definition is whether the entropy and sequence entropy of some shift are somehow related to each other. If we take the

sequence $\tau(i) = i$ for $i \in \mathbb{N}$ then the sequence entropy is the same as the topological one. However, some other choices of sequence τ may lead to different results. The Morse shift $X_{\mathcal{M}}$ is an example of space for which the values of the topological and sequence entropy are different. In [1] it is proved that the sequence entropy for the Morse shift $X_{\mathcal{M}}$ is equal to $\log 2$. However, the proof does not give the exact formula for the sequence which realizes the supremum from the definition. The main goal of this part of our paper is to show that the sequence $\tau(i) = 2^{2i} - 1$ works.

Theorem 16 *The sequence $\tau(i) = 2^{2i} - 1$ realizes the value of sequence entropy for Morse shift $X_{\mathcal{M}}$.*

Proof. Let us fix some $n \in \mathbb{N}$, $\epsilon < \frac{1}{2}$ and define the sequence $\tau(i) = 2^{2i} - 1$ for $i = 1, \ldots, n$. According to the definition we want to find a minimal (τ, ϵ, n)-spanning set for $X_{\mathcal{M}}$. If for some $x, y \in X_{\mathcal{M}}$ we have $x_{[0+\tau]} \neq y_{[0+\tau]}$ then there exists a position $j \in \{1, \ldots, n\}$ such that:

$$d(\sigma^{\tau(j)}(x), \sigma^{\tau(j)}(y)) > \frac{1}{2}.$$

Therefore any (τ, ϵ, n)-spanning set contains at least 2^n elements, in particular every element is a representant of a different pattern from 2^n possible patterns. Of course if $\epsilon = \frac{1}{2}$ the set W defined as above is the minimal (τ, ϵ, n)-spanning set for $X_{\mathcal{M}}$. Hence we have:

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log Span(\tau, \epsilon, n) =$$

$$\lim_{N \to \infty} \limsup_{n \to \infty} \frac{1}{n} \log Span(\tau, \frac{1}{2^N}, n) \geq \lim_{N \to \infty} \limsup_{n \to \infty} \frac{1}{n} \log 2^n = \log 2.$$

5. References

[1] Maass A., Shao S.; *Structure of Bounded Topological-Sequence-Entropy Minimal Systems*, Journal of the London Mathematical Society 76 (3), 2007, pp. 702–718.

[2] Kamae T., Zamboni L.; *Sequence Entropy and the Maximal Pattern Complexity of Infinite Words*, Ergodic Theory and Dynamical Systems 22 (4), 2002, pp. 1191–1199.

[3] Kamae T.; *Maximal Pattern Complexity as Topological Invariants*, preprint, Tokyo University, Available via http://www14.plala.or.jp/kamae/invariants.pdf.

[4] Restivo A., Salemi S.; *Overlap Free Words on Two Symbols*, Lecture Notes in Computer Science 192, Springer, New York 1985, pp. 198–206.

[5] Morse M., Hedlund G.A.; *Symbolic Dynamics*, American Journal of Mathematics 60(4), 1938, pp. 815–866.